

# Identification of signature genes of meta-clusters

LZ Liangtao Zheng SQ Shishang Qin XH Xueda Hu ZZ Zemin Zhang \*

Updated date: Jan 12, 2022

\*For correspondence: [zemin@pku.edu.cn](mailto:zemin@pku.edu.cn)

 An abbreviated version of this protocol was published in Science in Dec 2021

Pan-cancer single-cell landscape of tumor-infiltrating T cells

DOI: 10.1126/science.abe6474

## Detailed protocol

The signature genes of meta-clusters were identified based on a meta-analysis method. First, for each dataset, limma was used to identify differentially expressed genes by comparing cells from a meta-cluster to all cells from other meta-clusters. Second, moderated effect size and the variance of moderated effect size were estimated. Then the combined effect size, which was the weighted average of the effect sizes from all datasets, and corresponding *p*-values were calculated. *P*-values were adjusted using the Benjamini & Hochberg (BH) method implemented in the R function *p.adjust*. The signature genes of meta-clusters of pan-cancer were defined as those with a combined effect size larger than 0.15 (for genes that show significance in > 50% cancer types, 99.9% were above this value) and adjusted *p*-value less than 0.01. For cancer types with only one dataset, the significant genes were those with effect sizes > 0.15 and *p*-values < 0.01. For cancer types with more than one dataset, to combine the dataset level statistics to cancer-type level statistics, the above meta-analysis method was applied. Then if a gene showed significance in all datasets of the cancer type, the gene was significant; otherwise, the combined effect size and corresponding adjusted *p*-value were used for the discrimination, and threshold 0.15 and 0.01 were set for combined effect size and adjusted *p*-value respectively.

The code implemented the pipeline could be found in github (<https://github.com/Japrin/scPip>).

**How to cite:** (Readers should cite both the Bio-protocol preprint and the original research article where this protocol was used)

1. Zheng, L., Qin, S., Hu, X. and Zhang, Z. (2022). Identification of signature genes of meta-clusters. Bio-protocol Preprint. [bio-protocol.org/prep1501](https://bio-protocol.org/prep1501).
2. Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., Wu, N., Zhang, N., Zheng, H., Ouyang, H., Chen, K., Bu, Z., Hu, X., Ji, J. and Zhang, Z. (2021). Pan-cancer single-cell landscape of tumor-infiltrating T cells. Science 374(6574). DOI: [10.1126/science.abe6474](https://doi.org/10.1126/science.abe6474)

**Copyright:** Content may be subjected to copyright.